

GIST, 국제 AI 음향인식 대회서 심사위원상 수상

전기전자컴퓨터공학부 김홍국 교수 연구실

‘오니온에이아이(AunionAI)’ 팀 송윤아·이도현 학생

- IEEE 신호처리소사이어티(AASP) 주관 ‘DCASE 챌린지 2024’ 언어질의 기반 오디오 소스 분리(LASS) 부문에서 1위에 오른 데 이어 심사위원상까지 수상
- 거대언어모델(LLM) 기반 데이터 증강의 새로운 기법 제시해 성과... “연구실 성과가 사업화 가능성으로 이어진 사례로서, 보다 편리하고 안전한 삶 만드는 데 기여할 것”



▲ GIST 오디오지능연구실(AiTeR, 김홍국 교수)의 석박통합과정 송윤아 학생이 DCASE 챌린지 2024 워크숍에서 심사위원상을 수상하고 있다.

광주과학기술원(GIST, 총장 임기철)은 전기전자컴퓨터공학부 오디오지능연구실(AiTeR, 김홍국 교수) 학생들로 구성된 ‘GIST-AunionAI’팀(석박통합과정 송윤아·이도현)이 국제전기전자공학자협회(IEEE)의 신호처리 소사이어티(Audio and Acoustic Signal Processing(AASP))가 주관한 ‘국제 AI 음향 인식 경진대회(DCASE 챌린지 2024*)’ 워크숍에서 심사위원상(Judges’ award)을 수상했다고 밝혔다.

‘GIST-AunionAI’팀은 김홍국 교수가 창업한 오니온에이아이(AunionAI)의 지원을 받아 ‘DCASE 챌린지 2024’의 ‘언어질의 기반 오디오 소스 분리 과제(LASS: Language-Queried Audio Source Separation, Task 9)’ 부문에서 지난 7월 1위를 차지한 데 이어 10월에 열린 이번 워크숍에서는 국내외 우수한 연구기관을 제치고 ‘심사위원상’을 수상함으로써 연구의 혁신성과 기술의 완성도 및 논문의 품질 등 다양한 기준에서 우수한 성과를 인정받았다.

* **DCASE 챌린지**: 국제 음향 장면 및 이벤트 탐지 분류 경진대회(Challenge on Detection and Classification of Acoustic Scenes and Events)

‘GIST-AunionAI’팀이 이번에 받은 심사위원상은 DCASE 2024 Challenge에 제출된 기술 리포트를 기반으로 평가하여 **챌린지 순위와는 별개로 연구의 독창성과 우수성을 중심으로 검토해 과제(task)별 한 팀에만 수여된다.**

‘언어질의 기반 오디오 소스 분리(LASS) 기술’은 사용자가 입력한 텍스트에 따라 이에 맞는 오디오 신호를 분리하는 기술이다.

예를 들어 ‘누군가 무언가를 자르고 있으며, 바삭거리는 소리가 납니다’라는 텍스트 질의를 입력하면 주방의 다양한 소음이 혼재한 오디오에서 바삭거리며 자르는 소리만을 분리하여 추출하는 것이다. 이 기술은 언어와 오디오를 연결하는 생성형 AI 모델 개발의 기초를 제공하며, 자동 오디오 편집·멀티미디어 콘텐츠 검색·증강 청취 등 다양한 응용 분야에서 활용될 수 있다.

이번 워크숍에서 ‘GIST-AunionAI’ 팀은 다양한 오디오 지능을 표현할 수 있는 AI 기술들을 접목하여 고성능 언어질의 기반 오디오 소스 분리 기술을 발표했다. (논문명: Large-Language-Model-Based Caption Augmentation for Language-Queried Audio Source Separation)

연구팀은 ▲LLM(거대언어모델) 기반 프롬프트 기술 및 데이터 증강 기술 ▲사전 학습 훈련 모델*과 기존 모델의 추론 결과 융합 기술 그리고 ▲AI 능력 향상을 위한 앙상블 기술을 통해 AI 모델을 개선했다.

* 사전 학습 훈련 모델 : 대량의 데이터 세트에 대해 학습된 대용량의 모델



▲ GIST 오디오지능연구실(AiTÉR)의 (왼쪽부터) 김홍국 교수, 이도현 학생, 송윤아 학생

김홍국 교수는 "GIST 연구실과 (주)오니온에이아이가 협업하여 개발한 AI 모델이 **연구실에 머물지 않고 사업화 가능성을 제시했다**는 점에서 의의가 매우 크다"며 "특히, LLM 기반 오디오 생성 및 인식 AI 모델을 지속적으로 개선하는 노력과 함께 이를 다양한 분야에 적용하여 **보다 편리하고 안전한 삶을 위한 기술 발전에 기여**하겠다"고 말했다.

'GIST-AunionAI' 팀원들은 "김홍국 교수님의 지도와 아낌없는 지원 덕분에 국제경진 대회에서 좋은 성과를 얻을 수 있었다"고 말하며, "지금까지의 성과를 바탕으로 **오디오 인공지능 분야에서의 연구를 더욱 심화시켜 오디오지능 기술의 실용화와 사회적 가치 창출에 기여**할 계획"이라고 밝혔다.

한편, 음성 및 오디오와 관련된 다양한 AI 모델을 연구하고 있는 GIST 오디오지능 연구실(AiTeR, 지도교수 김홍국)은 언어질의 기반 음성 소스 분리뿐만 아니라 음향 이벤트 탐지, 음성 합성, 음성 잡음 제거, 음성 인식, 이상 상황 감지, 다국어 인식 및 번역 등 다양한 연구를 **국내 산업체와 대학, 연구기관을 비롯해 미국 매사추세츠공대(MIT) 등 해외 연구기관과 공동으로 진행하고 있다.**

'언어질의 기반 오디오 소스 분리'는 MIT 국제 공동 연구 사업과 GIST 과학기술혁신사업단의 '실용화 연구개발사업', 연구개발특구진흥재단의 '지역의 미래를 여는 과학기술 프로젝트' 사업의 지원으로 수행되었다.

DCASE 워크숍은 지난 **10월 23일부터 25일까지 일본 도쿄에서 진행**되었으며, 시상식은 **25일 '시나가와 시즌 테라스' 빌딩에서 열렸다.**